



Year: 2014

Genome-wide analysis of rare copy number variations reveals PARK2 as a candidate gene for attention-deficit/hyperactivity disorder

Jarick, Ivonne ; Volckmar, Anna-Lena ; Pütter, Carolin ; Pechlivanis, Sonali ; Nguyen, Trang T ; Dauvermann, Maria R ; Beck, Sebastian ; Albayrak, Özgür ; Scherag, Susann ; Gilsbach, Susanne ; Cichon, Sven ; Hoffmann, Per ; Degenhardt, Franziska ; Nöthen, Markus M ; Schreiber, Stefan ; Wichmann, H-Erich ; Jöckel, Karl-Heinz ; Heinrich, Joachim ; Tiesler, Carla M T ; Faraone, Stephen V ; Walitza, Susanne ; Sinzig, Judith ; Freitag, Christine ; Meyer, Jobst ; Herpertz-Dahlmann, Beate ; Lehmkuhl, Gerd ; Renner, Tobias J ; Warnke, Andreas ; Romanos, Marcel ; Lesch, Klaus-Peter ; Reif, Andreas ; Schimmelfmann, Benno G ; Hebebrand, Johannes ; Scherag, André ; Hinney, Anke

Abstract: Attention-deficit/hyperactivity disorder (ADHD) is a common, highly heritable neurodevelopmental disorder. Genetic loci have not yet been identified by genome-wide association studies. Rare copy number variations (CNVs), such as chromosomal deletions or duplications, have been implicated in ADHD and other neurodevelopmental disorders. To identify rare (frequency 1%) CNVs that increase the risk of ADHD, we performed a whole-genome CNV analysis based on 489 young ADHD patients and 1285 adult population-based controls and identified one significantly associated CNV region. In tests for a global burden of large (>500 kb) rare CNVs, we observed a nonsignificant ($P=0.271$) 1.126-fold enriched rate of subjects carrying at least one such CNV in the group of ADHD cases. Locus-specific tests of association were used to assess if there were more rare CNVs in cases compared with controls. Detected CNVs, which were significantly enriched in the ADHD group, were validated by quantitative (q)PCR. Findings were replicated in an independent sample of 386 young patients with ADHD and 781 young population-based healthy controls. We identified rare CNVs within the parkinson protein 2 gene (PARK2) with a significantly higher prevalence in ADHD patients than in controls ($P=2.8 \times 10^{-4}$) after empirical correction for genome-wide testing). In total, the PARK2 locus (chr 6: 162 659 756-162 767 019) harboured three deletions and nine duplications in the ADHD patients and two deletions and two duplications in the controls. By qPCR analysis, we validated 11 of the 12 CNVs in ADHD patients ($P=1.2 \times 10^{-3}$) after empirical correction for genome-wide testing). In the replication sample, CNVs at the PARK2 locus were found in four additional ADHD patients and one additional control ($P=4.3 \times 10^{-2}$). Our results suggest that copy number variants at the PARK2 locus contribute to the genetic susceptibility of ADHD. Mutations and CNVs in PARK2 are known to be associated with Parkinson disease.

DOI: <https://doi.org/10.1038/mp.2012.161>

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: <https://doi.org/10.5167/uzh-74491>
Journal Article
Published Version

Originally published at:

Jarick, Ivonne; Volckmar, Anna-Lena; Pütter, Carolin; Pechlivanis, Sonali; Nguyen, Trang T; Dauvermann, Maria R; Beck, Sebastian; Albayrak, Özgür; Scherag, Susann; Gilsbach, Susanne; Cichon, Sven; Hoffmann, Per; Degenhardt, Franziska; Nöthen, Markus M; Schreiber, Stefan; Wichmann, H-Erich; Jöckel, Karl-Heinz; Heinrich, Joachim; Tiesler, Carla M T; Faraone, Stephen V; Walitza, Susanne; Sinzig, Judith; Freitag, Christine; Meyer, Jobst; Herpertz-Dahlmann, Beate; Lehmkuhl, Gerd; Renner, Tobias J; Warnke, Andreas; Romanos, Marcel; Lesch, Klaus-Peter; Reif, Andreas; Schimmelmann, Benno G; Hebebrand, Johannes; Scherag, André; Hinney, Anke (2014). Genome-wide analysis of rare copy number variations reveals PARK2 as a candidate gene for attention-deficit/hyperactivity disorder. *Molecular Psychiatry*, 19(1):115-121.
DOI: <https://doi.org/10.1038/mp.2012.161>

ORIGINAL ARTICLE

Genome-wide analysis of rare copy number variations reveals *PARK2* as a candidate gene for attention-deficit/hyperactivity disorder

I Jarick¹, A-L Volckmar², C Pütter³, S Pechlivanis³, TT Nguyen¹, MR Dauvermann^{2,4}, S Beck², Ö Albayrak², S Scherag², S Gilsbach⁵, S Cichon^{6,7,8}, P Hoffmann^{7,8}, F Degenhardt^{7,8}, MM Nöthen^{7,8,9}, S Schreiber¹⁰, H-E Wichmann¹¹, K-H Jöckel³, J Heinrich¹¹, CMT Tiesler^{11,12}, SV Faraone¹³, S Walitza¹⁴, J Sinzig^{15,16}, C Freitag¹⁷, J Meyer¹⁸, B Herpertz-Dahlmann⁵, G Lehmkuhl¹⁵, TJ Renner¹⁹, A Warnke¹⁹, M Romanos^{19,20}, K-P Lesch^{21,22}, A Reif²³, BG Schimmelmann^{2,4}, J Hebebrand², A Scherag^{3,24} and A Hinney^{2,24}

Attention-deficit/hyperactivity disorder (ADHD) is a common, highly heritable neurodevelopmental disorder. Genetic loci have not yet been identified by genome-wide association studies. Rare copy number variations (CNVs), such as chromosomal deletions or duplications, have been implicated in ADHD and other neurodevelopmental disorders. To identify rare (frequency $\leq 1\%$) CNVs that increase the risk of ADHD, we performed a whole-genome CNV analysis based on 489 young ADHD patients and 1285 adult population-based controls and identified one significantly associated CNV region. In tests for a global burden of large (> 500 kb) rare CNVs, we observed a nonsignificant ($P = 0.271$) 1.126-fold enriched rate of subjects carrying at least one such CNV in the group of ADHD cases. Locus-specific tests of association were used to assess if there were more rare CNVs in cases compared with controls. Detected CNVs, which were significantly enriched in the ADHD group, were validated by quantitative (q)PCR. Findings were replicated in an independent sample of 386 young patients with ADHD and 781 young population-based healthy controls. We identified rare CNVs within the parkinson protein 2 gene (*PARK2*) with a significantly higher prevalence in ADHD patients than in controls ($P = 2.8 \times 10^{-4}$ after empirical correction for genome-wide testing). In total, the *PARK2* locus (chr 6: 162 659 756–162 767 019) harboured three deletions and nine duplications in the ADHD patients and two deletions and two duplications in the controls. By qPCR analysis, we validated 11 of the 12 CNVs in ADHD patients ($P = 1.2 \times 10^{-3}$ after empirical correction for genome-wide testing). In the replication sample, CNVs at the *PARK2* locus were found in four additional ADHD patients and one additional control ($P = 4.3 \times 10^{-2}$). Our results suggest that copy number variants at the *PARK2* locus contribute to the genetic susceptibility of ADHD. Mutations and CNVs in *PARK2* are known to be associated with Parkinson disease.

Molecular Psychiatry advance online publication, 20 November 2012; doi:10.1038/mp.2012.161

Keywords: ADHD; children; CNVs; GWAS; *PARK2*

INTRODUCTION

Attention-deficit/hyperactivity disorder (ADHD) represents one of the most common psychiatric disorders in children and adolescents with a worldwide prevalence rate of 5.2%.¹ According to the Diagnostic and Statistical Manual of Mental Disorders, Fourth

Edition (DSM-IV),² ADHD is characterized by pervasive and impairing symptoms of inattention, hyperactivity and increased impulsivity. Family, twin and adoption studies indicate that ADHD is a highly heritable disorder; heritability estimates are consistently around 0.8.^{3–7} However, neither genome-wide association studies

¹Institute of Medical Biometry and Epidemiology, University of Marburg, Marburg, Germany; ²Department of Child and Adolescent Psychiatry, University of Duisburg-Essen, Essen, Germany; ³Institute for Medical Informatics, Biometry and Epidemiology (IMIBE), University of Duisburg-Essen, Essen, Germany; ⁴University Hospital of Child and Adolescent Psychiatry, University of Bern, Bern, Switzerland; ⁵Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, RWTH Aachen University Clinics, Aachen, Germany; ⁶Institute of Neuroscience and Medicine (INM-1), Structural and Functional Organization of the Brain, Genomic Imaging, Research Center Juelich, Juelich, Germany; ⁷Institute of Human Genetics, University of Bonn, Bonn, Germany; ⁸Department of Genomics, Life and Brain Center, University of Bonn, Bonn, Germany; ⁹German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany; ¹⁰Institute of Clinical Molecular Biology, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany; ¹¹Institute of Epidemiology, German Research Center for Environmental Health, Helmholtz Center Munich, Neuherberg, Germany; ¹²Division of Metabolic Diseases and Nutritional Medicine, Dr von Hauner Children's Hospital, Ludwig-Maximilians-University of Munich, Munich, Germany; ¹³Departments of Psychiatry and of Neuroscience and Physiology, SUNY Upstate Medical University, Syracuse, NY, USA; ¹⁴Department of Child and Adolescent Psychiatry, University of Zurich, Zurich, Switzerland; ¹⁵Department for Child and Adolescent Psychiatry, University of Cologne, Cologne, Germany; ¹⁶Department for Child and Adolescent Psychiatry and Psychotherapy, LVR—clinic Bonn, Bonn, Germany; ¹⁷Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, JW Goethe-Universität Frankfurt am Main, Frankfurt am Main, Germany; ¹⁸Department of Neurobehavioral Genetics, Institute of Psychobiology, University of Trier, Trier, Germany; ¹⁹Department of Child and Adolescent Psychiatry, University of Wuerzburg, Wuerzburg, Germany; ²⁰Department of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy, University Hospital of Munich, Munich, Germany; ²¹Department of Psychiatry, Psychosomatics and Psychotherapy, Division of Molecular Psychiatry, ADHD Clinical Research Network, Laboratory of Translational Neuroscience, University of Wuerzburg, Wuerzburg, Germany; ²²Department of Neuroscience, School for Mental Health and Neuroscience, Maastricht University, Maastricht, The Netherlands and ²³Department of Psychiatry, Psychosomatics and Psychotherapy, University of Wuerzburg, Wuerzburg, Germany. Correspondence: Dr A Hinney, Department of Child and Adolescent Psychiatry, University of Duisburg-Essen, Virchowstraße 174, D-45147 Essen, Germany.

E-mail: anke.hinney@uni-due.de

²⁴These authors contributed equally to this work.

Received 14 December 2011; revised 21 September 2012; accepted 9 October 2012

(GWAS) nor large scale meta-analyses of GWAS has so far unequivocally identified specific genes conferring major risk (see Hinney *et al.*⁸).

Copy number variations (CNVs) are, by definition, chromosomal deletions or duplications of at least 1 kb up to several Mb that are variable in size among carriers. At a genome-wide level, thousands of CNVs have already been identified. Several CNVs were shown to contribute to neurodevelopmental disorders such as schizophrenia,⁹ Parkinson disease (PD)¹⁰ or autism.¹¹ Five genome-wide analyses of CNVs have been published in ADHD.^{12–16} Rare CNVs identified in ADHD patients were found to be preferentially located in neurodevelopmental genes,¹² that is, in genes reported as candidates in other neuropsychiatric disorders, such as autism, schizophrenia and Tourette syndrome, and in neurodevelopmental pathways. Large (>500 kb), rare ($\leq 1\%$ frequency) CNVs were found with an increased rate in British children with ADHD compared with controls from the British Birth Cohort. Locus-specific analyses revealed that this finding was largely based on several large CNVs in the chromosome 16p13.11 region; results for this region could be replicated in an Icelandic ADHD case–control sample.¹³ A second candidate region presumably harbouring CNVs associated with ADHD comprises the neuropeptide Y gene (*NPY*). The suggestive finding was based on a sample of 99 German children and adolescents with severe ADHD who were characterized by array comparative genomic hybridization.¹⁴ Recently, the detection of rare CNVs, which were present in Canadian ADHD cases but absent in a population-based control sample, identified new candidate ADHD susceptibility genes.¹⁵ A comparison of 896 ADHD cases with 2455 controls of European ancestry reported a significant excess of rare CNVs >100 kb among patients with ADHD.¹⁶ The strongest evidence for a single CNV locus was for duplications spanning the *CHRNA7* gene (15q13.3). This finding was replicated in additional 2242 ADHD cases and 8552 controls. In addition, an overlap between nominally associated CNV and single nucleotide polymorphism (SNP) data was detected for 13 enriched biological pathways (including cholesterol-related and CNS developmental pathways), and for 1 specific candidate at the gene level: *CHRNA7*.¹⁷ Elia *et al.*¹⁸ reported CNV analyses for 2493 ADHD cases and 9222 controls of European ancestry. They found an enrichment of CNVs in metabotropic glutamate receptor genes in patients with ADHD.

We followed-up the mentioned previous findings by investigating the hypothesis of ‘common disease—many rare variants’¹⁹ in our study. In more detail, we investigated the role of rare CNVs in ADHD by performing a genome-wide CNV association study in 489 patients with ADHD and 1285 population-based control individuals based on Illumina SNP arrays (Illumina Ltd, Little Chesterford, UK; ADHD patients: Human660W-QuadV1; controls: HumanHap550v3 BeadArrays). We searched for regions where ADHD patients had an increased number of CNVs (deletions and insertions) compared with controls. Replication of our findings was conducted in an independent sample of 386 patients with ADHD and 781 healthy controls on the basis of genome-wide Affymetrix SNP array data (Affymetrix Genome-Wide Human SNP Array 5.0, Affymetrix Ltd, High Wycombe, UK).

MATERIALS AND METHODS

Participants

GWAS sample. In total, 504 patients recruited for the GWAS step (see Hinney *et al.*⁸) were assessed for the diagnosis ADHD according to DSM-IV.² The corresponding subtypes and basic characteristics of those 489 ADHD patients whose molecular data met our pre- and post-calling quality control (QC) criteria (see Supplementary Text) are given in Table 1. The GWAS cases are all German minors (age range: 6–18 years, mean age: 11.0 ± 2.7 years) with ADHD, recruited and phenotypically characterized in six psychiatric outpatient units for children and adolescents (Aachen,

Cologne, Essen, Marburg, Regensburg and Würzburg). The ascertainment strategy and inclusion criteria have been described previously.^{20–22}

In addition, controls not screened for ADHD were drawn from three German population-based epidemiological studies in adults (for details see Cichon *et al.*²³) for the GWAS step: (a) the Heinz Nixdorf RECALL (Risk Factors, Evaluation of Coronary Calcification, and Lifestyle) study²⁴ (HNR, $n = 383$), (b) PopGen²⁵ ($n = 490$) and (c) KORA²⁶ ($n = 488$). The recruitment areas were Essen, Bochum and Mülheim (Ruhr area) for (a), Schleswig-Holstein (Northern Germany) for (b) and Augsburg (Southern Germany) for (c), respectively. Molecular data from 1285 of the 1361 controls (see Table 1) met our pre- and post-calling QC criteria. Compared with the ADHD cases, the population-based controls were less frequently male (cases: 81.0% male; controls: 50.7% male) and older (age range cases: 6–18 years; age range controls: 25–75 years).

Replication sample. The case group of the replication sample consisted of 461 young (age range: 6–19 years, mean age: 10.4 ± 2.4 years) patients of German ancestry with ADHD, who were recruited and phenotypically characterized in two outpatient clinics at the Departments of Child and Adolescent Psychiatry, Psychosomatics and Psychotherapy of the Universities of Homburg and Würzburg, Germany. Patients were included if they were diagnosed with ADHD according to DSM-IV², subtypes and basic characteristics of the 386 analysed (after pre- and post-calling QC) ADHD cases are given in Table 1. The ascertainment strategy and inclusion criteria have been described previously.^{22,27} Core descriptive statistics of the ADHD cases of the replication sample were comparable to those of the ADHD GWAS sample (83.7% male; age range: 6–19 years).

For the replication sample, 1063 controls were chosen from two ongoing German population-based prospective birth cohorts: the influence of life-style factors on the Immune System and Allergies Plus environment and genetics (LISAPlus)²⁸ study and the German Infant study on the influence of Nutrition Intervention Plus environment and genetics (GINIplus).²⁹ Briefly, the two birth cohorts consist of healthy full-term newborns, who were recruited between September 1955 and January 1999 in Munich, Wesel, Leipzig and Bad Honnef and followed-up to age 10. A detailed description of screening and recruitment has been provided elsewhere.^{28,29} Any probands without questionnaire information ($n = 111$) or who were categorized as above the normal range at the age of 10 on the Strengths and Difficulties Questionnaire on the scales for hyperactivity/inattention or with the Strengths and Difficulties Questionnaire total difficulties score ($n = 118$) were excluded from the analysis ($n = 229$). A summary of the 781 analysed (after pre- and post-calling QC) replication controls is provided in Table 1.

Ethnicity in both the GWAS and the replication sample was assigned to patients and controls according to self-reported ancestry (all German). Written informed consent was given by all individuals or by their parents in case of minors. The study protocols were approved by the respective Institutional Review Board or Ethics Committees and they were conducted in accordance with The Declaration of Helsinki.

Statistical analysis

Details on genotyping, the procedures for CNV detection and QC for CNV calls are given in the Supplementary Text. All association analyses were performed on the QC filtered, rare CNV calls were investigated using the PLINK software (version 1.07).³⁰ As primary analyses, we tested the hypothesis that particular rare CNVs might be found at an increased frequency in ADHD cases compared with controls. Locus-specific tests of association were performed (one-sided χ^2 tests) and significance was assessed via permutation (empirical P -values based on 100 000 permutations) at a pointwise as well as at a genome-wide level. In more detail, we calculated the frequency of rare CNVs in ADHD patients, and we compared it to the frequency in the controls. The frequency was calculated at each unique start and stop site for rare CNVs that met all of the defined QC measures (defined in the Supplementary Text). Each site (5047 sites in total, located in 1083 non-overlapping genomic CNV containing regions) was assessed for a difference in CNV frequency between groups with the use of a permutation-based Fisher's exact test in PLINK. We refer to a locus in the sense of a susceptibility locus as a genomic region that is exclusively made up of adjacently tested sites for which significantly more rare CNVs were observed in ADHD cases than in controls. These analyses were undertaken for all rare CNVs as well as stratified according to CNV type, that is, deletion or duplication. In the GWAS sample, empirical, genome-wide corrected P -values were generated by permuting affection status and simultaneously preserving the correlation structure of CNVs (100 000

Table 1. Descriptive data for ADHD cases and controls following QC of the GWAS and the replication step

	GWAS ADHD cases	GWAS controls ^a	Replication ADHD cases	Replication controls
Total, <i>n</i>	489 ^b	1285 ^c	386 ^d	781 ^e
Sex				
Male, <i>n</i> (%)	396 (81.0%)	651 (50.7%)	323 (83.7%)	402 (51.5%)
Female, <i>n</i> (%)	93 (19.0%)	634 (49.3%)	63 (16.3%)	379 (48.5%)
Age, mean (s.d.; years)	11.0 (2.7)	—	10.5 (2.5)	10.0 (0.2)
ADHD subtype ^f				
Combined type, <i>n</i> (%)	360 (73.6%)	—	316 (81.9%)	0 (0%) ^g
Predominantly inattentive type, <i>n</i> (%)	105 (21.5%)	—	55 (14.2%)	0 (0%) ^g
Predominantly hyperactive-impulsive type, <i>n</i> (%)	24 (4.9%)	—	15 (3.9%)	0 (0%) ^g

Abbreviations: ADHD, attention-deficit/hyperactivity disorder; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition; GWAS, genome-wide association study; QC, quality control.

^aThe GWAS controls were not screened. Their age ranges from 25 to 75 years. ADHD subtypes are unknown.

^bA total of 504 ADHD patients were recruited for the GWAS step, after QC 489 ADHD patients were left for full analysis.

^cA total of 1361 controls were recruited for the GWAS step, after QC 1285 probands were left for full analysis.

^dA total of 461 ADHD patients were recruited for the replication step, after QC 386 ADHD patients were left for full analysis.

^eA total of 1063 controls were recruited for the replication step, after QC 781 probands were left for full analysis.

^fCurrent DSM-IV² diagnosis.

^gForty-four of the 781 controls for the full analysis of the replication step were categorized as borderline on the SDQ hyperactivity/inattention scale.

permutations) to simulate the null hypothesis of no association. In other words, we used the permutation resampling method to correct for the multiple testing problem, which occurs when testing any identified locus of rare CNVs.

By application of PLINK's case-control 'cnv-enrichment-test' function, we additionally tested whether CNVs in the *PARK2* gene are enriched in ADHD GWAS cases compared with GWAS controls. In contrast to Fisher's exact test, the 'cnv-enrichment-test' is robust to case-control differences in CNV size or CNV rate.³¹ Enrichment in cases is reported as one-sided empirical *P*-value using 100 000 permutations.

Analysis of the replication sample was performed to confirm the finding of the GWAS sample at the *PARK2* locus (Results section) and we focussed the testing on this single rare CNV locus for an association to ADHD. Consequently, pointwise empirical *P*-values (100 000 permutations) for the replication sample were not corrected for multiple testing. The statistical analyses in the GWAS sample were repeated for quantitative (q)PCR-validated CNVs at the *PARK2* locus as part of the sensitivity analysis. We applied a significance level α of 0.05 (globally for the genome-wide testing and locally for the replication sample).

As secondary sensitivity analyses (see Supplementary Text), we also assessed the genome-wide frequency of CNVs in ADHD cases compared with controls according to the average number of CNVs per sample. We expected more CNVs in the ADHD cases based on the literature.³² Thus, one-sided tests were applied to all rare CNVs, as well as to rare deletions and duplications only; and genome-wide multiple testing was dealt with using 100 000 permutations. Finally, we likewise tested whether CNVs in the ADHD cases were larger in size than those in the control group based on the average size of CNVs per individual.

CNV validation and replication analysis at the *PARK2* locus

We performed real-time qPCR experiments to validate the CNVs by a Duplex TaqMan CNV assay (Applied Biosystems, Darmstadt, Germany, assay Hs03615859_cn at chr6: 162 696 987 ± 50 bp, NCBI36/hg18) as described previously.³³ Individual copy number status was determined for each ADHD patient of the GWAS sample. Briefly, every PCR was performed as a triplicate for each individual of the GWAS ADHD cases and the results from the qPCR were analysed using the software CopyCaller 1.0 (Applied Biosystems). In cases, 11 of the 12 CNVs identified with PennCNV covering the *PARK2* locus (Results section) were technically validated by qPCR. qPCR experiments did not reveal any further CNV carrier, which was undetected in previous SNP array-based CNV detection analyses. Thus, CNVs at the *PARK2* locus of GWAS ADHD patients could be validated with both, a low false-positive and a low false-negative rate. For a subset of controls (HNR controls), CNV calls, which were estimated to cover the *PARK2* gene, were validated by qPCR. Apart from one potential CNV carrier, who was incorporated into statistical analysis at the *PARK2* locus, we additionally considered the five HNR controls for which CNVs were estimated to flank the *PARK2* locus (*n* = 3) or for which CNVs were called but excluded due to their small size (spanned < 15 probes) in the course of

our CNV QC procedure (*n* = 2). Moreover, we additionally included six randomly chosen control subjects of the HNR control sample. CNV analyses were performed blinded to the likely CNV status of the controls. For the *PARK2* locus, all analysed CNV states could be validated. For the KORA and PopGen controls, no DNA was available for qPCR validation. However, given the high technical validation rate in the available DNA samples, validity of CNV calls was presumed to be comparably high for KORA and PopGen controls. Although false-negative and false-positive rates are unknown for the GWAS controls group, there is no obvious reason to expect that these rates would significantly differ between cases and controls. Notably, the low frequency of *PARK2* CNVs in control subjects was consistently reported in the 'Database of Genomic Variants' (<http://projects.tcag.ca/variation>) and in two previous publications, where CNVs at the *PARK2* gene were absent in 2026 healthy, population-based controls¹² and in 1409 healthy children,¹¹ respectively. In a conservative manner, we excluded the non-validated CNV from the ADHD group, whereas all four CNVs in the control group were included in all analyses.

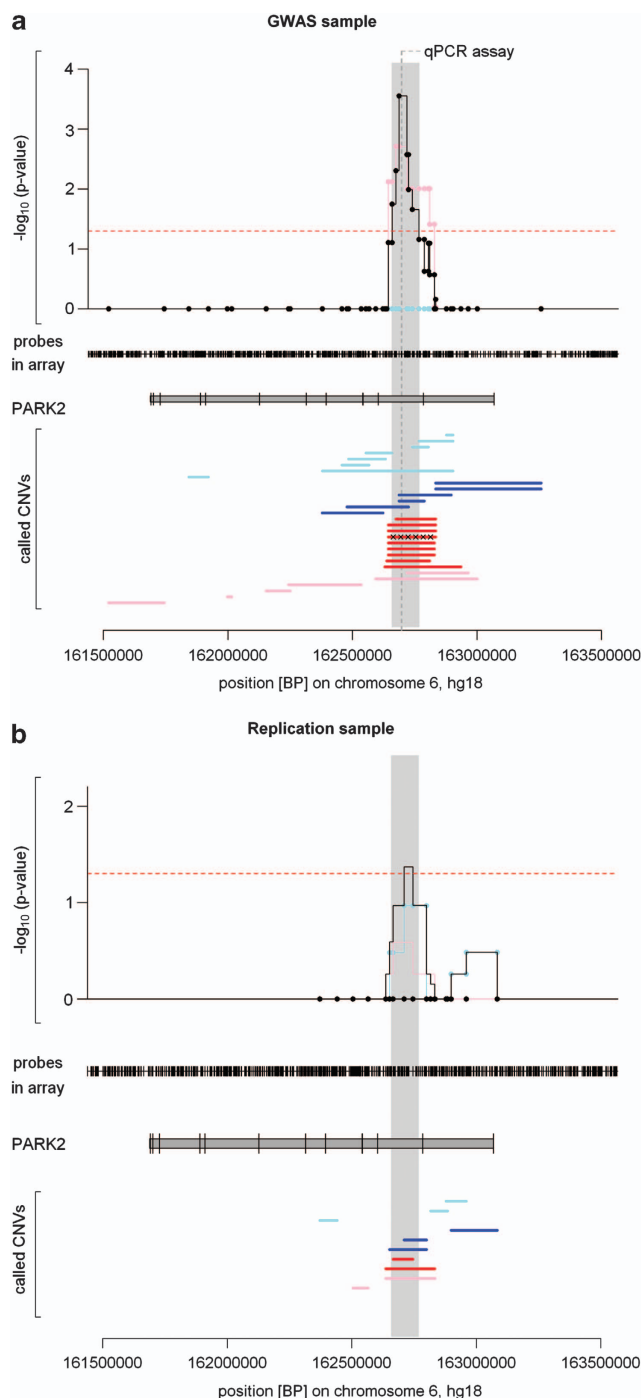
We tested the CNV association with ADHD in the *PARK2* gene in an independent replication sample. Details on genotyping, CNV calling and QC are described in the Supplementary Text.

RESULTS

The GWAS sample included 489 ADHD cases and 1285 controls (Table 1) with high-quality SNP array data for full CNV analysis. Comparison of the CNV sets identified in the ADHD patients and in the controls showed no increased overall frequency of CNVs in ADHD cases (Supplementary Text). After exclusion of common (frequency > 1%) CNVs, 2432 rare CNVs (592 in ADHD cases; 1840 in controls) with an increased length in ADHD cases (average CNV size: 226.3 kb (range: 9.3–2830.8 kb) in ADHD cases; 186.4 kb (range: 5.6–4479.6 kb) in controls) were included in the association analysis. Although there was a difference in the sex distribution between ADHD patients (81.0% males) and control subjects (50.7% males), there was no evidence for significant difference in the rate at which rare CNVs were called in males compared with females in either cases or comparison subjects (data not presented). All rare CNVs > 500 kb are listed in Supplementary Table S1.

With regard to previous observations,^{13,16,17} we first looked at our data with respect to a potential overall enrichment of rare CNVs in ADHD cases compared with controls (Supplementary Table S7) and in terms of an enrichment for loci implicated in neurodevelopmental disorders, such as autism or schizophrenia (data not shown). There was no evidence for an increased burden of rare CNVs in ADHD patients (*P* = 0.997). We additionally performed comparative analyses on rare CNVs stratified by their size. Interestingly, with increasing size thresholds, we observed a

stronger trend of association between large, rare CNVs and ADHD, which is in accordance with the reports of previous studies.^{13,16} Despite the fact that none of the comparisons resulted in a nominally significance (that is, $P < 0.05$): there was a 1.126-fold enriched rate ($P = 0.271$) and a 1.133-fold higher proportion ($P = 0.253$) of subjects carrying at least one rare CNV > 500 kb in the ADHD cohort. The rate of rare CNVs > 500 kb observed in ADHD cases was 10.4%, which is similar to the rates of 12.2 and 12.5% reported in previous studies.^{13,16} Limiting our analysis to rare CNVs > 2 Mb, we observed a 3.065-fold enrichment ($P = 0.074$) in ADHD cases relative to control subjects. However, due to the potential bias in individual CNV rate and average size, which differentiates cases and controls in our GWAS sample (see Supplementary Text), we did not follow-up these data.



Differences between distributions in cases and controls may rather result from different technical genotyping procedures, than indicating association effects.

Locus-specific association tests for an overrepresentation of CNVs, including both deletions and duplications, in ADHD cases in comparison with controls revealed only one genome-wide significant genomic region within the *PARK2* gene with a P -value of 2.8×10^{-4} empirically corrected for genome-wide testing (Figure 1, Supplementary Table S8). This locus for which we observed more rare CNVs in ADHD cases than in controls is located at chr6: 162 659 756–162 767 019 (NCBI36/hg18). We refer to this region as the *PARK2* locus. In total, this locus is covered by 12 CNVs among the ADHD patients (2.45%, three deletions (0.61%) and nine duplications (1.84%)) and four CNVs among the controls (0.31%, two deletions (0.16%) and two duplications (0.16%)), all these CNVs extend into the coding region (either exon 2 or exon 3) of *PARK2*. Locus-specific association tests, stratified according to CNV type (deletion or duplication), did not reveal further genomic regions with genome-wide significant results. The *PARK2* locus alone showed a genome-wide significant enrichment of duplications in ADHD cases compared with controls ($P = 1.9 \times 10^{-3}$ after empirical correction for genome-wide testing, Figure 1). In contrast, we did not observe a genome-wide significant excess of deletions for any of the tested regions harbouring rare CNVs.

Enrichment of CNVs at the *PARK2* gene for GWAS ADHD cases compared with GWAS controls was also supported by the robust 'cnv-enrichment-test'³¹ (one-sided empirical $P = 8.8 \times 10^{-4}$).

Copy number status at the *PARK2* locus of all ADHD patients in the GWAS sample was reevaluated by qPCR analyses (Figure 1). Apart from one duplication, each CNV status was technically validated. Even after reanalysis of all rare CNVs with exclusion of the non-validated duplication, the finding for the *PARK2* locus remained genome-wide significant (empirically corrected $P = 1.2 \times 10^{-3}$, Supplementary Figure S7). We observed no differences in ADHD subtypes or basic characteristics like age, sex and intelligence quotient (IQ) values between carriers of CNVs at the *PARK2* locus and the 489 ADHD patients of the GWAS sample (Supplementary Table S2).

Next, we assessed an independent sample of 386 ADHD patients and 781 healthy controls to replicate our finding of an excess of CNVs in ADHD patients at the *PARK2* locus. We replicated the excess ($P = 4.3 \times 10^{-2}$, Figure 1) of CNVs in the ADHD replication sample ($n = 4$ (1.04%), two duplications (0.52%) and

Figure 1. Results for the *PARK2* locus in the GWAS and in the replication sample. Each panel consists of four parts: (called CNVs, probes analysed and association tests): CNVs: red (pink) bars represent duplications in an ADHD case (control), blue (lightblue) bars indicate a deletion of an ADHD case (control). *PARK2* gene: the marks indicate the coding regions (NCBI36/hg18). Association tests: permutation-based one-sided $-\log_{10}$ -transformed P -values for association tests; the black (pink; lightblue) line represents association tests for an increased frequency of segmental CNV data independent of type (deletions; duplications) in cases compared with controls. The significance level $P = 0.05$ is highlighted as a dashed red line. The chromosomal region offering genome-wide significantly more CNVs in ADHD patients than in controls is highlighted by grey vertical shading. (a) Results for the GWAS sample. The presented P -values are genome-wide empirically corrected. The chromosomal region covered by the qPCR assay used for validation of PennCNV's CNV calls is shown as a darkgrey vertical dashed line within this region of genome-wide significance. The duplication that could not be validated by qPCR analysis is marked by 'x'. Results of association tests after exclusion of the non-validated case duplication are shown in Supplementary Figure S7. (b) Results for the replication sample. The presented P -values are pointwise corrected.

two deletions (0.52%) compared with replication controls (one duplication (0.13%)). Similar to the initial finding, four of the five CNVs in the *PARK2* locus extend into the coding exon 2 of the gene. Owing to the small number of CNVs we did not stratify by CNV status.

Finally, in addition to our GWAS discovery analysis, we assessed ADHD CNV candidate loci from previous reports^{12–18} in our GWAS sample using a hypothesis driven approach. First, we did not identify any putative CNV carrier at the *NPY* locus.¹⁴ Second, we found a total of eight CNVs (four in cases (one duplication and three deletions), four in controls (one duplication and three deletions); Supplementary Figure S8) at the chr 16p13.11 locus.¹³ Six of the eight CNVs were > 500 kb (four in cases (one duplication and three deletions), two in controls (both deletions); uncorrected $P = 0.053$ (0.13 for deletions, 0.28 for duplications)). Thus, our data partly support the relevance of CNVs of the chromosome 16p13.11 locus for ADHD. Third, we determined the number of CNVs identified in our GWAS sample that overlapped with any of the 28 reported CNVs as well as with the corresponding 22 candidate genes reported by Elia et al.¹² (Supplementary Tables S3 and 4). In addition to *PARK2*, we reidentified two further genes, *CHL1* and *PTPRD*, with a higher frequency of CNVs in ADHD patients compared with controls. We identified two CNVs (one deletion, one duplication) in ADHD patients (patients frequency (F_{patients}) = 0.41%) and one deletion in the controls (controls frequency (F_{controls}) = 0.08%) that overlapped *CHL1*. The set of CNVs involving *PTPRD* included three deletions in ADHD patients (F_{patients} = 0.61%) and five CNVs (four deletions and one duplication) in controls (F_{controls} = 0.39%). Fourth, we did not identify any CNVs at *GRM5*, *GRM8*, *GRM7* or *GRM1*.¹⁸ Finally, we compared the frequency of CNVs identified in our GWAS ADHD cases and controls that overlapped with any of 22 rare CNV loci exclusively found in Canadian ADHD patients and not in Canadian or German population-based controls¹⁵ (Supplementary Tables S5 and 6). For two loci, at chr 6p24.2 (covering *GCNT2*) and at chr 16p11.2, we reidentified one CNV exclusively in ADHD cases (F_{patients} = 0.20%), respectively. In addition, we observed two further loci with higher frequencies of CNVs in ADHD cases than in controls. We found four duplications (F_{patients} = 0.82%) in ADHD patients and seven duplications (F_{controls} = 0.54%) in the controls at the chr 7q36.3 locus (covering *PTPRN2*); as well as five duplications in ADHD cases (F_{patients} = 1.02%) and eight CNVs (one deletion, seven duplications) in the controls (F_{controls} = 0.62%, deletions: F_{controls} = 0.08%, duplications: F_{controls} = 0.54%) at the chr 15q13 locus.^{16,17}

DISCUSSION

Within two moderately sized ADHD case–control samples, we showed that children with ADHD have a significantly increased frequency of CNVs, including both deletions and duplications, at the *PARK2* locus compared with controls. The initial impression based on the observations in the GWAS sample, that this excess of CNVs at the *PARK2* locus is mainly driven by duplications, could not be substantiated in the replication sample. In conclusion, our data indicate that the *PARK2* locus influences susceptibility to ADHD when it comprises either deletions or duplications. Similarly, deletions and duplications on chromosome 16p11.2 were previously observed to be collectively associated with autism.³⁴ Moreover, CNVs at chromosome 1q21.1 were shown to be associated with congenital heart defects, developmental delay, schizophrenia and related psychoses,³⁵ whereas more recently microcephaly and macrocephaly were consistently observed in patients with 1q21.1 deletions and duplications, respectively, along with a spectrum of developmental delay, neuropsychiatric abnormalities, dysmorphic features and congenital anomalies.³⁵ Regarding CNVs at the *PARK2* locus, their role in the genetics of ADHD needs to be further investigated in more detail.

Interestingly, the *PARK2* gene, which was previously associated with schizophrenia,³⁶ was one of the neurodevelopmental genes

that were identified to harbour CNVs in ADHD cases but not in healthy controls in the first published genome-wide investigation of CNVs in ADHD.¹² However, in the study by Elia et al. CNVs in the *PARK2* gene did not achieve genome-wide significance—two CNVs were found in 335 ADHD cases and none in 2026 healthy controls.

The precise function of the *PARK2* gene, which encodes parkin, is unknown; however, the encoded protein is a component of a multiprotein E3 ubiquitin ligase complex that mediates the targeting of substrate proteins for proteasomal degradation. Parkin was suggested to increase dopamine uptake by enhancing the ubiquitination and degradation of misfolded dopamine transporter, so as to prevent it from interfering with the oligomerization and cell surface expression of native dopamine transporter.³⁷ With this function, parkin would enhance the precision of dopaminergic transmission, increase the efficiency of dopamine utilization and reduce dopamine toxicity on neighbouring cells.³⁷ Mutations in the *PARK2* gene have been reported as a cause of PD³⁸ and autosomal recessive juvenile PD.³⁸ Alternative splicing of this gene produces multiple transcript variants encoding distinct isoforms. Additional splice variants of this gene have been described but currently lack transcript support.³⁹ CNVs within or surrounding genes involved in the ubiquitin pathway, including among others *PARK2*, were observed to be significantly enriched in cases affected by autism spectrum disorders compared with a healthy control cohort.¹¹ More recently, *PARK2* CNVs (including deletions and duplications) were also shown to be associated with PD susceptibility.¹⁰ Fitting to these results is an observation, we have described earlier in a clinical retrospective study including patients with early and late onset PD that found association of PD with symptoms of ADHD in childhood.⁴⁰ In addition, the *PARK2* gene, including rare inherited copy number changes at this locus, was also implicated in a study pertaining to schizophrenia.³⁶

CNVs at the *PARK2* locus were shown to be significantly more common in children with ADHD but were not restricted to this group. Furthermore, among the ADHD patients of the GWAS sample, carriers of CNVs covering the *PARK2* locus showed no differences from non-carriers by ADHD subtype, age, sex or IQ values. This observation is in concordance with recent results that showed a general lack of differences in clinical or cognitive characteristics between children with ADHD, with and without large rare CNVs.⁴¹ The only reported significant difference was a higher rate of intellectual disability (IQ < 70) among CNV carriers compared with non-carriers, whereas most CNV carriers did not have an intellectual disability. As an IQ > 70 was an inclusion criteria for our study, we could not analyse CNV burden according to IQ value and thus unfortunately could not compare our results with the mentioned study.⁴¹

Apart from the *PARK2* locus, our observations support findings of previous studies.^{12,13,15} In parts, we substantiated the relevance of CNVs at the chromosome 16p13.11 locus for ADHD. The consensus region of large CNVs at chromosome 16p13.11 spans seven genes: *MPV17L*, *C16orf45*, *KIAA0430*, *NDE1*, *MYH11*, *C16orf63* and *ABCC1*. Of particular interest is the nuclear distribution gene E homologue 1 (*NDE1*), which regulates neuronal activity and interacts with the *disrupted in schizophrenia 1* gene, a risk factor for schizophrenia and related illnesses.⁴² CNVs spanning this gene region have also been associated with intellectual disability,⁴³ autism⁴³ and schizophrenia.⁴⁴ We also found an excess of CNVs, and particular of duplications, in the 15q13.3 region spanning *CHRNA7*, which was the strongest finding from the Williams et al.¹⁶ ADHD CNV study. Moreover, our data support evidence for their potential role in ADHD for CNVs covering four genes additional to *PARK2*: *CHL1*, *PTPRD*, *GCNT2* and *PTPRN2*. Previously, CNVs spanning these genes were exclusively found in ADHD patients.^{12,15} In our study, CNVs covering the genes *PARK2*, *CHL1*, *PTPRD* and *PTPRN2* were observed at a higher frequency in ADHD patients compared with controls. A deletion covering *GCNT2* was exclusively found in one ADHD patient. Missense

mutations in the *CHL1* gene were identified to be associated with schizophrenia.⁴⁵ Then, SNPs in *PTPRD* were identified as risk factors for the restless legs syndrome⁴⁶ and disruptions of the *GCNT2* gene have previously been reported in autism spectrum disorders.⁴⁷ Finally, with regard to previous reports of ADHD being a frequent phenotypic component in patients with deletions or duplications at 15q13 and 16p11.2,^{48,49} our results underline the role of rare CNVs at these two loci for ADHD. However, the investigation of loci previously implicated in ADHD in our sample is limited by statistical power. The modest *P*-values we obtained at each locus are almost certainly the result of this limited power to detect rare variants.

With regard to reported results of previous studies,^{13,16} we additionally tested for a global burden of rare CNVs in ADHD cases relative to control subjects. Although not reaching significance, we consistently observed a trend for an overall enrichment of large (> 500 kb), rare CNVs harboured by ADHD cases compared with controls. However, as CNVs were not validated at the genome-wide level, these results need to be interpreted with caution.

As our study is based on custom genotyping chips, it remains to be determined whether a sequencing approach or the use of higher-resolution platforms would detect a broader range of rare CNVs associated with ADHD. In addition, much larger sample sizes may yield sufficient power to detect common genetic variants with modest or marginal effects. With respect to the use of different SNP arrays, we limited our analysis to the intersecting set of SNPs between chip platforms and applied conservative quality-control criteria. Nevertheless, our observation that rare CNVs in ADHD patients were on average larger but less frequent than those in controls (see Supplementary Text) may in part be explained by the inability of a precise determination of CNV breakpoints on the basis of SNP chip data. To support our main finding, however, the association of *PARK2* CNVs with ADHD was validated by qPCR and replicated in an independent sample.

In summary, our results support the role of structural variants at the *PARK2* locus for ADHD genetics. Moreover, our data support the further investigation of CNVs involving neurodevelopmental genes, such as *CHL1*, *PTPRD*, *GCNT2*, *PTPRN2* and *NDE1*, as well as deletions and duplications at the 15q13 and 16p11.2 regions for ADHD genetics.

CONFLICT OF INTEREST

KH Jöckel and A Scherag are responsible statisticians in several clinical trials and in this position they receive support money from various pharmaceutical companies, among them MEDICE Arzneimittel Pütter GmbH and KG. In the past year, Dr Faraone received consulting income and research support from Shire and Alcobra and research support from the National Institutes of Health. In previous years, he received consulting fees or was on Advisory Boards or participated in continuing medical education programs sponsored by: Shire, McNeil, Janssen, Novartis, Pfizer and Eli Lilly. Dr Faraone receives royalties from books published by Guilford Press: *Straight Talk about Your Child's Mental Health* and Oxford University Press: *Schizophrenia: The Facts*. S Walitza receives industry funding research from Vifor Pharma, Switzerland and was on the speakers' bureau of Eli Lilly, Janssen-Cilag and Astra Zeneca. SW receives research funding from DFG, SNF, BMFFSJ and FP7. CM Freitag was on the speakers' bureau of Eli Lilly and Novartis, and was an expert consultant for Desitin. Dr Herpertz-Dahlmann receives industry research funding from Vifor Pharma, Schweiz. In previous years, she was on the Advisory Board of Eli Lilly and received sponsoring by AstraZeneca, Eli Lilly, Novartis and Janssen Cilag. She received research support from the Federal Ministry of Education and Research and the German Research Society (Deutsche Forschungsgemeinschaft). G Lehmkuhl receives industry research funding from Lilly Deutschland GmbH, Germany. A Warnke receives industry research funding from Shire and in previous years from Eli Lilly. B Schimmelfmann is a member of the speakers' bureau of Eli Lilly, BMS, Janssen and Novartis. The other authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank the children and their families for their participation and support to this study. We are also grateful to all probands from the community-based cohorts of

PopGen, KOR, those from the Heinz Nixdorf RECALL (HNR) study, and the GINIplus and LISAPlus cohorts. We thank the Heinz Nixdorf Foundation, Germany, for the generous support of the HNR study. We thank the German Research Association (DFG) who funded the GWAS analyses and confirmatory studies (He1446/9-1 to J Hebebrand, KP Lesch, A Hinney and T Renner, KFO 125, SFB 581, SFB TRR 58/A5, GRK 1253 to KP Lesch; ME 1923/5-1, ME 1923/5-3 to J Meyer and CM Freitag, GRK 1389 to J Meyer, SCHA 542/10-3 to H Schäfer) and the Bundesministerium für Bildung und Forschung (BMBF 01GV0605 to KP Lesch). We thank the START-Program EK 119/05 of the Medical Faculty, RWTH Aachen, Germany. The European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 245009 supported this study.

REFERENCES

- Polanczyk G, de Lima MS, Horta BL, Biederman J, Rohde LA. The worldwide prevalence of ADHD: a systematic review and meta-regression analysis. *Am J Psychiatry* 2007; **164**: 942–948.
- American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Diseases (DSM-IV)*. 4th ed. American Psychiatric Publishing: Washington, DC, 1994.
- Heiser P, Friedel S, Dempfle A, Konrad K, Smidt J, Grabarkiewicz J et al. Molecular genetic aspects of attention-deficit/hyperactivity disorder. *Neurosci Biobehav Rev* 2004; **28**: 625–641.
- Faraone SV, Perlis RH, Doyle AE, Smoller JW, Goralnick JJ, Holmgren MA et al. Molecular genetics of attention deficit/hyperactivity disorder. *Biol Psychiatry* 2005; **57**: 1313–1323.
- Faraone SV, Mick E. Molecular genetics of attention deficit hyperactivity disorder. *Psychiatr Clin North Am* 2010; **33**: 159–180.
- Franke B, Neale BM, Faraone SV. Genome-wide association studies in ADHD. *Hum Genet* 2009; **126**: 13–50.
- Freitag CM, Rhode LA, Lempp T, Romanos M. Phenotypic and measurement influences on heritability estimates in childhood ADHD. *Eur Child Adolesc Psychiatry* 2010; **19**: 311–323.
- Hinney A, Scherag A, Jarick I, Albayrak Ö, Pütter C, Pechlivanis S et al. Genome-wide association study in German patients with attention deficit/hyperactivity disorder. *Am J Med Genet B Neuropsychiatr Genet* 2011; **156**: 888–897.
- Stefansson H, Rujescu D, Cichon S, Pietiläinen OP, Ingason A, Steinberg S et al. Large recurrent microdeletions associated with schizophrenia. *Nature* 2008; **455**: 232–236.
- Pankratz N, Dumitriu A, Hetrick KN, Sun M, Latourelle JC, Wilk JB et al. Copy number variation in familial parkinson disease. *PLoS One* 2011; **6**: e20988.
- Glessner JT, Wang K, Cai G, Korvatska O, Kim CE, Wood S et al. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. *Nature* 2009; **459**: 569–573.
- Elia J, Gai X, Xie HM, Perin JC, Geiger E, Glessner JT et al. Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. *Mol Psychiatry* 2010; **15**: 637–646.
- Williams NM, Zaharieva I, Martin A, Langley K, Mantripragada K, Fossdal R et al. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: a genome-wide analysis. *Lancet* 2010; **376**: 1401–1408.
- Lesch KP, Selch S, Renner TJ, Jacob C, Nguyen TT, Hahn T et al. Genome-wide copy number variation analysis in attention-deficit/hyperactivity disorder: association with neuropeptide Y gene dosage in an extended pedigree. *Mol Psychiatry* 2011; **16**: 491–503.
- Lionel AC, Crosbie J, Barbosa N, Goodale T, Thiruvahindrapuram B, Rickaby J et al. Rare copy number variation discovery and cross-disorder comparisons identify risk genes for ADHD. *Sci Transl Med* 2011; **3**: 95ra75.
- Williams NM, Franke B, Mick E, Richard JL, Freitag CM, Gill M et al. Genome-wide analysis of copy number variants in attention deficit hyperactivity disorder: the role of rare variants and duplications at 15q13.3. *Am J Psychiatry* 2012; **169**: 195–204.
- Stergiakouli E, Hamshere M, Holmans P, Langley K, Zaharieva I, deCODE Genetics et al. Investigating the contribution of common genetic variants to the risk and pathogenesis of ADHD. *Am J Psychiatry* 2012; **169**: 186–194.
- Elia J, Glessner JT, Wang K, Takahashi N, Shtir CJ, Hadley D et al. Genome-wide copy number variation study associates metabotropic glutamate receptor gene networks with attention deficit hyperactivity disorder. *Nat Genet* 2011; **44**: 78–84.
- Mayo O. The rise and fall of the common disease-common variant (CD-CV) hypothesis: how the sickle cell disease paradigm led us all astray (or did it?). *Twin Res Hum Genet* 2007; **10**: 793–804.
- Hebebrand J, Dempfle A, Saar K, Thiele H, Herpertz-Dahlmann B, Linder M et al. A genome-wide scan for attention-deficit/hyperactivity disorder in 155 German sib-pairs. *Mol Psychiatry* 2006; **11**: 196–205.
- Schimmelfmann BG, Friedel S, Dempfle A, Warnke A, Lesch KP, Walitza S et al. No evidence for preferential transmission of common valine allele of the Val66Met

- polymorphism of the brain-derived neurotrophic factor gene (BDNF) in ADHD. *J Neural Transm* 2007; **114**: 523–526.
- 22 Romanos M, Freitag C, Jacob C, Craig DW, Dempfle A, Nguyen TT *et al*. Genome-wide linkage analysis of ADHD using high-density SNP arrays: novel loci at 5q13.1 and 14q12. *Mol Psychiatry* 2008; **13**: 522–530.
 - 23 Cichon S, Mühleisen TW, Degenhardt FA, Mattheisen M, Miró X, Strohmaier J *et al*. Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am J Hum Genet* 2011; **88**: 372–381.
 - 24 Schmermund A, Möhlenkamp S, Stang A, Grönemeyer D, Seibel R, Hirche H *et al*. Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Risk factors, evaluation of coronary calcium and lifestyle. *Am Heart J* 2002; **144**: 212–218.
 - 25 Krawczak M, Nikolaus S, von Eberstein H, Croucher PJ, El Mokhtari NE, Schreiber S. PopGen: population-based recruitment of patients and controls for the analysis of complex genotype-phenotype relationships. *Community Genet* 2006; **9**: 55–61.
 - 26 Wichmann HE, Gieger C, Illig T. Group, MONICA/KORA Study. KORA-gen—resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* 2005; **67**: S26–S30.
 - 27 Renner TJ, Walitza S, Dempfle A, Eckert L, Romanos M, Gerlach M *et al*. Allelic variants of SNAP25 in a family-based sample of ADHD. *J Neural Transm* 2008; **115**: 317–321.
 - 28 Zutavern A, Brockow I, Schaaf B, Bolte G, von Berg A, Diez U *et al*. Timing of solid food introduction in relation to atopic dermatitis and atopic sensitization: results from a prospective birth cohort study. *Pediatrics* 2006; **117**: 401–411.
 - 29 Berg A, Kramer U, Link E, Bollrath C, Heinrich J, Brockow I *et al*. Impact of early feeding on childhood eczema: development after nutritional intervention compared with the natural course—the GINIplus study up to the age of 6 years. *Clin Exp Allergy* 2010; **40**: 627–636.
 - 30 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al*. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
 - 31 Raychaudhuri S, Korn JM, McCarroll SA, International Schizophrenia Consortium, Altshuler D, Sklar P *et al*. Accurately assessing the risk of schizophrenia conferred by rare copy-number variation affecting genes with brain function. *PLoS Genet* 2010; **6**: pii: e1001097.
 - 32 Girirajan S, Brkanac Z, Coe BP, Baker C, Vives L, Vu TH *et al*. Relative burden of large CNVs on a range of neurodevelopmental phenotypes. *PLoS Genet* 2011; **7**: e1002334.
 - 33 deKovel CG, Trucks H, Helbig I, Mefford HC, Baker C, Leu C *et al*. Recurrent microdeletions at 15q11.2 and 16p13.11 predispose to idiopathic generalized epilepsies. *Brain* 2010; **133**: 23–32.
 - 34 Weiss LA, Shen Y, Korn JM, Arking DE, Miller DT, Fossdal R *et al*. Association between microdeletion and microduplication at 16p11.2 and autism. *N Engl J Med* 2008; **358**: 667–675.
 - 35 Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, Bacino CA, Sahoo T *et al*. Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat Genet* 2008; **40**: 1466–1471.
 - 36 Xu B, Roos JL, Levy S, van Rensburg EJ, Gogos JA, Karayiorgou M. Strong association of de novo copy number mutations with sporadic schizophrenia. *Nat Genet* 2008; **40**: 880–885.
 - 37 Jiang H, Jiang Q, Feng J. Parkin increases dopamine uptake by enhancing the cell surface expression of dopamine transporter. *J Biol Chem* 2004; **279**: 54380–54386.
 - 38 Crosiers D, Theuns J, Cras P, Van Broeckhoven C. Parkinson disease: insights in clinical, genetic and pathological features of monogenic disease subtypes. *J Chem Neuroanat* 2011; **42**: 131–141.
 - 39 Chien HF, Rohé CF, Costa MD, Breedveld GJ, Oostra BA, Barbosa ER *et al*. Early-onset Parkinson's disease caused by a novel parkin mutation in a genetic isolate from north-eastern Brazil. *Neurogenetics* 2006; **7**: 13–19.
 - 40 Walitza S, Melfsen S, Herhaus G, Scheuerpflug P, Warnke A, Müller T *et al*. Association of Parkinson's disease with symptoms of attention deficit hyperactivity disorder in childhood. *J Neural Transm* 2007; **72**(Suppl): 311–315.
 - 41 Langley K, Martin J, Agha SS, Davies C, Stergiakouli E, Holmans P *et al*. Clinical and cognitive characteristics of children with attention-deficit hyperactivity disorder, with and without copy number variants. *Br J Psychiatry* 2011; **199**: 398–403.
 - 42 Bradshaw NJ, Christie S, Soares DC, Carlyle BC, Porteous DJ, Millar JK. NDE1 and NDEL1: multimerisation, alternate splicing and DISC1 interaction. *Neurosci Lett* 2009; **449**: 228–233.
 - 43 Ullmann R, Turner G, Kirchhoff M, Chen W, Tonge B, Rosenberg C *et al*. Array CGH identifies reciprocal 16p13.1 duplications and deletions that predispose to autism and/or mental retardation. *Hum Mutat* 2007; **28**: 674–682.
 - 44 Kirov G, Grozeva D, Norton N, Ivanov D, Mantripragada KK, Holmans P *et al*. Support for the involvement of large copy number variants in the pathogenesis of schizophrenia. *Hum Mol Genet* 2009; **18**: 1497–1503.
 - 45 Sakurai K, Migita O, Toru M, Arinami T. An association between a missense polymorphism in the close homologue of L1 (CHL1, CALL) gene and schizophrenia. *Mol Psychiatry* 2002; **7**: 412–415.
 - 46 Yang Q, Li L, Yang R, Shen GQ, Chen Q, Foldvary-Schaefer N *et al*. Family-based and population-based association studies validate PTPRD as a risk factor for restless legs syndrome. *Mov Disord* 2011; **26**: 516–519.
 - 47 van der Zwaag B, Franke L, Poot M, Hochstenbach R, Spierenburg HA, Vorstman JA *et al*. Gene-network analysis identifies susceptibility genes related to glyco-biology in autism. *PLoS One* 2009; **4**: e5324.
 - 48 Shinawi M, Liu P, Kang SH, Shen J, Belmont JW, Scott DA *et al*. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. *J Med Genet* 2010; **47**: 332–341.
 - 49 Miller DT, Shen Y, Weiss LA, Korn J, Anselm I, Bridgemohan C *et al*. Microdeletion/duplication at 15q13.2q13.3 among individuals with features of autism and other neuropsychiatric disorders. *J Med Genet* 2009; **46**: 242–248.



This work is licensed under the Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

Supplementary Information accompanies the paper on the Molecular Psychiatry website (<http://www.nature.com/mp>)